

THE CHARACTERISTIC IMSET POLYTOPE FOR DIAGNOSIS MODELS

JING XI AND RURIKO YOSHIDA

ABSTRACT. In 2010, M. Studený, R. Hemmecke, and Linder explored a new algebraic description of graphical models, characteristic imsets. Compare with standard imsets, characteristic imsets have several advantages: they are still unique vector representative of conditional independence structures, they are 0-1 vectors, and they are more intuitive in terms of graphs than standard imsets. After defining characteristic imset polytope as the convex hull of all characteristic imsets for a given set of nodes, they also showed that a model selection in graphical models, which essentially is a problem of maximizing a quality criterion, can be converted into an integer programming problem on the characteristic imset polytope. However, this integer programming problem is very hard in general. Therefore, here we focus on diagnosis models which can be described by Bipartite graphs with a set of m nodes and a set of n nodes for any $m, n \in \mathbb{Z}_+$, and their characteristic imset polytope. In this paper, first, we will show that the characteristic imsets for diagnosis models have very nice properties including that the number of non-zero coordinates is at most is $n \cdot (2^m - 1)$, and with these properties we are able to find a combinatorial description of all edges of the characteristic imset polytopes for diagnosis models. Then we prove that these characteristic imset polytopes are direct products of n many $(2^m - 1)$ dimensional simplices. Finally, we end the paper with further questions in this topic.

Keywords: graphical model, characteristic imset polytope, diagnosis model, Bipartite graph

1. INTRODUCTION

Bayesian networks (BNs), also known as belief networks, Bayes networks, Bayes(ian) models or probabilistic directed acyclic graphical models, find their applications to model knowledge in many areas, such as computational biology and bioinformatics (gene regulatory networks, protein structure, gene expression analysis [3] learning epistasis from GWAS data sets [4]) and medicine [13]. BNs are a part of the family of probabilistic graphical models (GMs). These graphical structures represent knowledge about probabilistic structures for a statistical model. More precisely, each node in the graph represents a random variable and an edge between the nodes represents probabilistic dependencies among the random variables corresponding to the nodes adjacent to the edge [6]. BNs correspond to GM structure known as a directed acyclic graph (DAG) defined by the set of nodes (vertices) and the set of directed edges.

In order to infer parameters from the observed data set, we first apply a model selection criterion called *quality criterion*, which provides a way to construct highly predictive BN models from data by choosing the graph which gives the given criteria, such as Bayesian Information Criteria (BIC) [8] or Akaike Information Criteria (AIC) [1], maximum (see [10] for more details on quality criterions). Intuitively a quality criterion is a function, $Q(G, D)$, which takes a DAG, G , and an observed data set, D , to evaluate how good the DAG G to explain the observed data D . Note that different DAGs, G_1, G_2 may have the same conditional independences (CIs). In that case we say G_1, G_2 are *Markov equivalent*. When researchers wish to infer the CIs of the BN structure from the observed data set one represents each set of Markov equivalent graphs by one graph called the *essential graph* the corresponding Markov equivalence class of DAGs [2]. In this paper we focus on quality criterions $Q(G, D)$, such that $Q(G_1, D) = Q(G_2, D)$ if and only if G_1, G_2 are Markov equivalent.

Since in general there are super exponentially many essential graphs with a fixed set of nodes N , maximizing the quality criterion, $\mathcal{Q}(G, D)$, over all possible essential graphs with N is known to be NP-hard. Studený developed an algebraic representation of each essential graph G called a *standard imset*, of G , which is an integral vector representation of G in $\mathbb{R}^{2^{|N|}-|N|-1}$. From the view of this setting a criterion function $\mathcal{Q}(G, D)$ is a dot product of vectors in $\mathbb{R}^{2^{|N|}-|N|-1}$. In 2010, M. Studený, J. Vomlel, and R. Hemmecke showed that maximizing the $\mathcal{Q}(G, D)$ over all essential graphs can be formulated as a linear programming problem over the convex hull of standard imsets for all possible essential graphs [12]. This gives us a systematic way to find the best criterion with the optimality certificate rather than finding the best criterion by the brute-force search. Then M. Studený, R. Hemmecke, and Linder explored an alternative vector representative of the BN structure, called *characteristic imsets*. Compare with standard imsets, characteristic imsets have several advantages: they are still unique vector representative of conditional independence structures; they are 0-1 vectors; and they are more intuitive in terms of graphs than standard imsets [11].

In general, however, the dimension of the convex hull of the characteristic imsets with the fixed set of nodes N , called *characteristic imset polytope*, is exponentially large and there are double exponentially many vertices as well as facets of the characteristic imset polytope. Thus it is infeasible to optimize by software if $|N| > 6$. In order to solve the LP problem for a larger $|N|$, we need to understand the structure of the characteristic imset polytope, such as combinatorial description of edges and facets of the polytope so that we might be able to apply a simplex method to find an optimal solution. However, in general, it is challenging because there are too many facets and too many edges of the polytope. Therefore here we focus on a particular family of BN models, namely *diagnosis models*.

In medical studies, researchers are often interested in probabilistic models in order for them to correctly diagnose a disease from a patient symptoms. The diagnoses models, also known as the Quick Medical Reference (QMR) diagnostic model, is introduced in [9] to diagnose a disease from a given set of symptoms of a patient. Therefore, here we focus on diagnosis models (e.g., [7]). Under this model, a DAG representing the model is a bipartite graph with two sets of nodes, one representing m diseases and one representing n symptoms, and set of directed edges from nodes representing diseases to nodes representing symptoms (see Figure 1 for an example).

In this paper, first, we will show that the dimension of the characteristic imset polytope for diagnosis models with m diseases and n symptoms is $n \cdot (2^m - 1)$ which is much smaller than $2^{(m+n)} - (n + m) - 1$. Second, we are able to find an explicit combinatorial description of all edges of the characteristic imset polytopes for diagnosis models with fixed m and n , that is, if G_1, G_2 are graphs representing two diagnosis models such that all symptoms have the same parents in G_1 and in G_2 except one symptom, then the characteristic imsets representing G_1, G_2 form an edge of the characteristic imset polytope for diagnosis models. Then we prove that these characteristic imset polytopes are direct products of n many $(2^m - 1)$ dimensional simplices.

This paper is organized as follows. In Section 2 we introduce notation and state some definitions. Section 3 shows propositions and their proofs, and Section 4 shows our main results. Then we will show some examples in Section 5 and discuss our future work in Section 6.

2. NOTATION AND DEFINITIONS

In this section we state some notation and remind readers some definitions.

Definition 2.1. A Diagnosis Model can be described by a Bipartite Graph whose nodes $N = \{a_1, \dots, a_m\} \cup \{b_1, \dots, b_n\}$ can be divided into disjoint sets $A = \{a_1, \dots, a_m\}$ and $B = \{b_1, \dots, b_n\}$. Nodes in A can be interpreted as diseases and nodes in B can be interpreted as symptoms. Every single edge can only be drawn from a disease to a symptom. An example is

given by Figure 1.

Define notation: $\mathcal{G}_{m,n} = \{\text{All possible Bipartite graphs defined in Definition 2.1 for fixed } m \text{ and } n\}$.

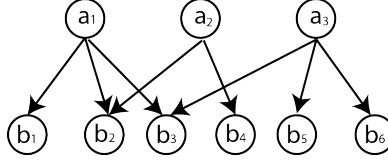


FIGURE 1. An example of Bipartite Graph, $m = 3$, $n = 6$.

3. PROPOSITIONS AND PROOFS

Recall that we have the definition of Characteristic Imset.

Definition 3.1. Let G be an acyclic directed graph over N . The **characteristic imset** for G can be introduced as a zero-one vector c_G with components $c_G(S)$ where $S \subseteq N$, $|S| \geq 2$ given by

$$c_G(S) = 1 \iff \exists i \in S \text{ such that } j \in pa_G(i) \text{ for } \forall j \in S \setminus \{i\}$$

where $j \in pa_G(i)$ means G includes the edge from j to i .

Proposition 3.2. Assume $|N| > 2$ and G is a Bipartite graph. $A = \{a_1, \dots, a_m\}$ and $B = \{b_1, \dots, b_n\}$ are defined in Definition 2.1. Then $c_G(T)$ is possible to take value 1 if and only if T has the form of $a_{i_1} \dots a_{i_k} b_j$, where $1 \leq k \leq m$, $\{i_1, \dots, i_k\} \subseteq \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$.

Proof. Notice that for $\forall T \subseteq N$, $|T| \geq 2$, we can write T in the following form:

$$(3.1) \quad T = a_{i_1} \dots a_{i_k} b_{j_1} \dots b_{j_l}, \text{ where } \begin{aligned} 0 \leq k \leq m, \{i_1, \dots, i_k\} &\subseteq \{1, \dots, m\}, \\ 0 \leq l \leq n, \{j_1, \dots, j_l\} &\subseteq \{1, \dots, n\}, \\ k + l &\geq 2. \end{aligned}$$

Now we need to prove that l can neither be 0 nor greater than 1.

- (a) If $l = 0$. For $\forall s, t \in \{i_1, \dots, i_k\}$, by Definition 2.1, there cannot be an edge from a_s to a_t . This means $a_s \notin pa_G(a_t)$. Hence $\forall t \in \{i_1, \dots, i_k\}$, $T \setminus \{a_t\} \not\subseteq pa_G(a_t)$. $c_G(T) = 0$.
- (b) If $l > 1$. Similar with above, by Definition 2.1, for $\forall s', t' \in \{j_1, \dots, j_l\}$, $b_{s'} \notin pa_G(b_{t'})$. Moreover, for $\forall t \in \{i_1, \dots, i_k\}$ and $t' \in \{j_1, \dots, j_l\}$, $b_{t'} \notin pa_G(a_t)$. $c_G(T) = 0$.

□

Proposition 3.3. Notation same as Proposition 3.2. Suppose T has the form of $a_{i_1} \dots a_{i_k} b_j$, where $1 \leq k \leq m$, $\{i_1, \dots, i_k\} \subseteq \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$, then $c_G(T) = \prod_{s=i_1, \dots, i_k} c_G(a_s b_j)$.

Proof. As shown in part (a) in the proof of Proposition 3.2, for $\forall s, t \in \{i_1, \dots, i_k\}$, $a_s \notin pa_G(a_t)$. Therefore:

$$\begin{aligned} c_G(T) &= 1 \text{ if and only if } \{a_{i_1} \dots a_{i_k}\} \subseteq pa_G(b_j); \\ &\text{If and only if } a_s \in pa_G(b_j), \text{ for } \forall s = i_1, \dots, i_k; \\ &\text{If and only if } c_G(a_s b_j) = 1, \text{ for } \forall s = i_1, \dots, i_k. \end{aligned}$$

Recall that $c_G(T)$ is binary. Thus $c_G(T) = \prod_{s=i_1, \dots, i_k} c_G(a_s b_j)$.

□

Remark 3.4. Another way to see Proposition 3.3 is that for a diagnosis model, the whole characteristic imset is determined by only $m \cdot n$ coordinates, $c_G(a_i b_j)$, $i \in \{1, \dots, m\}$, $j \in \{1, \dots, n\}$, which can also be interpreted as the existence of edge point from a_i to b_j . From this remark, it is straightforward to get Proposition 3.6. For a related example, see Example 5.1.

Remark 3.5. Proposition 3.3 also implies that $\forall G \in \mathcal{G}_{m,n}$, G can be determined by $pa_G(b_j)$, $b_j \in B$, and $pa_G(b_j)$, $b_j \in B$ are completely irrelevant.

Proposition 3.6. *Notation same as Definition 2.1. Fix m and n . The number of elements in $\mathcal{G}_{m,n}$ is 2^{mn} .*

Proof. Consider a bipartite graph $G \in \mathcal{G}_{m,n}$. At most $m \cdot n$ possible edges can be assigned from nodes in a set A to nodes in a set B . Note that there are

$$\sum_{k=0}^{mn} \binom{mn}{k} = 2^{mn}$$

many ways to assign edges from nodes in a set A to nodes in a set B . \square

Proposition 3.7. *Notation same as Definition 2.1. Fix m and n . If we consider the characteristic imset for an arbitrary Bipartite graph in $\mathcal{G}_{m,n}$, the number of non-zero coordinates is at most $n \cdot (2^m - 1)$.*

Proof. Consider T by classify its cardinality.

- When $2 \leq |T| \leq m + 1$, by Proposition 3.2, the coordinate related to T can be 1 if and only if T has the form of $a_{i_1} \dots a_{i_{|T|-1}} b_j$, $\{i_1, \dots, i_{|T|-1}\} \subseteq \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$.

Fix $|T|$, the number of this kind of T is: $\binom{m}{|T|-1} \cdot n$.

- When $|T| > m + 1$, $\exists b_{j_1}, b_{j_2} \in \{1, \dots, n\}$ s.t. $b_{j_1}, b_{j_2} \in T$. By Proposition 3.2, $c_G(T) \equiv 0$.

Therefore, the number of possible non-zero coordinates is:

$$\sum_{|T|=2}^{m+1} \binom{m}{|T|-1} \cdot n = n \cdot \sum_{k=1}^m \binom{m}{k} = n \cdot (2^m - 1).$$

\square

Remark 3.8.

- Notation same as Definition 2.1. For a fixed N , by Proposition 3.2 and 3.7, we can define $\mathcal{S}_{m,n}$ as the support of $\{c_G : G \in \mathcal{G}_{m,n}\}$. We know that:

$$\mathcal{S}_{m,n} = \{T : \exists G \in \mathcal{G}_{m,n} \text{ such that } c_G(T) = 1\} \subset \mathcal{P}(N)$$

where $\mathcal{P}(N)$ is the power set of N .

- Recall that in elementary geometry,
 - a **closed convex polyhedron** (which will be indicated as **polyhedron** for short) in \mathbb{R}^q can be defined by a system of linear inequalities:

$$\{\mathbf{x} \in \mathbb{R}^q : A\mathbf{x} \leq \mathbf{b}\}$$
 where A is a $p \times q$ matrix in $\mathbb{R}^{p \times q}$ and \mathbf{b} is a $p \times 1$ vector in $\mathbb{R}^{p \times 1}$;
 - a **closed convex polytope** (which will be indicated as **polytope** for short) is defined as the convex hull of a finite set of points;
 - if a polyhedron is bounded, then it is a polytope.
- A **d-simplex** is a d-dimensional polytope which has exactly $d + 1$ vertices. It is notated as Δ_d .
- For more details on polyhedral geometry see [14].

- Notice that Proposition 3.7 also implies: for fixed m and n , the dimension of the polytope of characteristic imsets for all elements in $\mathcal{G}_{m,n}$ (we call it **characteristic imset polytope** in the following and use $\mathbf{P}_{m,n}$ as the notation) is at most $n \cdot (2^m - 1)$. We will prove that it is actually exactly $n \cdot (2^m - 1)$.
- Because characteristic imsets are all 0-1 vectors, it is obvious that for fixed m and n , the set of vertices of characteristic imset polytope is exactly $\{c_G : G \in \mathcal{G}_{m,n}\}$.

4. THEOREMS AND PROOFS

Theorem 4.1. *Notation same as Definition 2.1. Fix m and n . The dimension of the characteristic imset polytope is exactly $n \cdot (2^m - 1)$.*

Proof. For an N , we can consider the standard basis \mathbf{e}_T , $T \subset N$, which have the same coordinates with characteristic imsets and $\mathbf{e}_T(T_0) = 1$ if and only if $T_0 = T$, $\forall T_0 \subset N$.

It is obvious that 1.) $\{c_G, G \in \mathcal{G}_{m,n}\} \subset \mathbb{R}^{2^{m+n}-(m+n+1)}$, 2.) $\{\mathbf{e}_T, T \in \mathcal{S}_{m,n}\}$ is a basis of $\mathbb{R}^{n \cdot (2^m - 1)}$ which is embedded in $\mathbb{R}^{2^{m+n}-(m+n+1)}$ and 3.) $\{c_G, G \in \mathcal{G}_{m,n}\}$ can be written as a linear combination of $\{\mathbf{e}_T, T \in \mathcal{S}_{m,n}\}$. We are going to prove that $\{\mathbf{e}_T, T \in \mathcal{S}_{m,n}\}$ can be expressed as a linear combination of $\{c_G, G \in \mathcal{G}_{m,n}\}$.

Notice that $\{\mathbf{e}_T, T \in \mathcal{S}_{m,n}\}$ is equivalent with $\{\mathbf{e}_T, T \subset N \text{ and } T \text{ has the form of } a_{i_1} \dots a_{i_k} b_j, \text{ where } 1 \leq k \leq m, \{i_1, \dots, i_k\} \subseteq \{1, \dots, m\} \text{ and } j \in \{1, \dots, n\}\}$, we can prove the theorem by induction.

- When $k = 1$, i.e. $T = a_i b_j$ where $a_i \in A$ and $b_j \in B$, we can choose graph G which has only one edge $(a_i b_j)$. Then $c_G = \mathbf{e}_T$.
- Consider $T_k = a_{i_1} \dots a_{i_k} b_j$ where $1 < k \leq m$, $\{i_1, \dots, i_k\} \subseteq \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$. Suppose for $\forall T$, $|T| \leq k$, \mathbf{e}_T can be written as a linear combination of $\{c_G, G \in \mathcal{G}_{m,n}\}$.

Let G be the Bipartite graph with k edges: $(a_{i_l} b_j)$, $l = 1 \dots k$. Then:

$$\mathbf{e}_{T_k} = c_G - \sum_{T_a \subset \{a_{i_1}, \dots, a_{i_k}\}, 0 < |T_a| < k} \mathbf{e}_{T_a \cup \{b_j\}}$$

By assumption, $\mathbf{e}_{T_a \cup \{b_j\}}$ can be expressed as a linear combination of $\{c_G, G \in \mathcal{G}_{m,n}\}$, for $\forall T_a \subset \{a_{i_1}, \dots, a_{i_k}\}, 0 < |T_a| < k$, therefore, \mathbf{e}_{T_k} can be written as a linear combination of $\{c_G, G \in \mathcal{G}_{m,n}\}$.

□

Remark 4.2. Notice that for the special case $n = 1$, Theorem 4.1 and Proposition 3.6 claim that $\mathbf{P}_{m,1}$ has 2^m vertices and the dimension of $\mathbf{P}_{m,n}$ is $n \cdot (2^m - 1)$. This directly lead to Theorem 4.3.

Theorem 4.3. *Fix m , $\mathbf{P}_{m,1}$ is a simplex with dimension $2^m - 1$, i.e. $\mathbf{P}_{m,1} = \Delta_{2^m - 1}$.*

4.1. Combinatorial Description of Edges on $\mathbf{P}_{m,n}$.

Definition 4.4. Graphs $G, H \in \mathcal{G}_{m,n}$ are called **neighbors** if c_G and c_H form an edge in the characteristic imset polytope.

Lemma 4.5. *Notation same as Definition 2.1. Fix m , then for arbitrary two distinct graphs, $G_1, G_2 \in \mathcal{G}_{m,1}$, G_1 and G_2 are neighbors, i.e. c_{G_1} and c_{G_2} form an edge in the characteristic imset polytope.*

Proof. Let $N = \{a_1, \dots, a_m, b_1\}$. We are going to prove \exists a cost vector w , such that $w \cdot c_{G_1} = w \cdot c_{G_2} > w \cdot c_{G_3}$ for $\forall G_3 \in \mathcal{G}_{m,n}$ distinct with G_1 and G_2 .

By Remark 3.5, G_1 and G_2 are determined by $pa_{G_1}(b_1)$ and $pa_{G_2}(b_1)$, respectively. Let's discuss by two scenarios of $pa_{G_1}(b_1)$ and $pa_{G_2}(b_1)$.

- (1) One is a subset of the other. WLOG, suppose $pa_{G_1}(b_1) \subsetneq pa_{G_2}(b_1)$.

Define notations: $A_1 = pa_{G_1}(b_1)$, $A_2 = pa_{G_2}(b_1)$, $A_{2 \setminus 1} = pa_{G_2}(b_1) \setminus pa_{G_1}(b_1)$, and $A_{comp} = (pa_{G_2}(b_1))^c$. Notice that $A_{2 \setminus 1} \neq \emptyset$, A_1 and A_{comp} can be \emptyset . A_1 , $A_{2 \setminus 1}$ and A_{comp} is a partition of N .

Consider a cost vector w whose coordinates match the coordinates of characteristic im-sets.

- If $|A_{2 \setminus 1}| > 1$, define w as:

$$w(T) = \begin{cases} c & \text{for } T = a_i b_j, a_i \in A_1 \\ -c & \text{for } T = a_i b_j, a_i \notin A_1 \\ |A_{2 \setminus 1}| \cdot c & \text{for } T = A_{2 \setminus 1} \cup \{b_1\} \\ 0 & \text{for } T \subset N, |T| > 2, \text{ and } T \neq A_{2 \setminus 1} \cup \{b_1\} \end{cases}$$

where c is a positive number.

Then for $\forall G_3 \in \mathcal{G}_{m,n}$, we have:

$$\begin{aligned} w \cdot c_{G_3} &= |A_1 \cap pa_{G_3}(b_1)| \cdot c - |pa_{G_3}(b_1) \setminus A_1| \cdot c + |A_{2 \setminus 1}| \cdot c \cdot c_{G_3}(A_{2 \setminus 1} \cup \{b_1\}) \\ &= |A_1 \cap pa_{G_3}(b_1)| \cdot c - |pa_{G_3}(b_1) \cap A_{2 \setminus 1}| \cdot c - |pa_{G_3}(b_1) \cap A_{comp}| \cdot c + |A_{2 \setminus 1}| \cdot c \cdot c_{G_3}(A_{2 \setminus 1} \cup \{b_1\}) \end{aligned}$$

Now notice that in the above equation:

- * $|A_1 \cap pa_{G_3}(b_1)| \cdot c \leq |A_1| \cdot c$, where “=” holds if and only if $A_1 \subset pa_{G_3}(b_1)$;
- * $-|pa_{G_3}(b_1) \cap A_{2 \setminus 1}| \cdot c + |A_{2 \setminus 1}| \cdot c \cdot c_{G_3}(A_{2 \setminus 1} \cup \{b_1\}) \leq 0$, where “=” holds if and only if $pa_{G_3}(b_1) \cap A_{2 \setminus 1} = \emptyset$ or $A_{2 \setminus 1}$;
- * $-|pa_{G_3}(b_1) \cap A_{comp}| \cdot c \leq 0$, where “=” holds if and only if $pa_{G_3}(b_1) \cap A_{comp} = \emptyset$.

Therefore, $w \cdot c_{G_3} \leq |A_1| \cdot c$ where “=” holds if and only if $G_3 = G_1$ or G_2 .

- If $|A_{2 \setminus 1}| = 1$, let $A_{2 \setminus 1} = \{a_q\}$, define w as:

$$w(T) = \begin{cases} c & \text{for } T = a_i b_j, a_i \in A_1 \\ -c & \text{for } T = a_i b_j, a_i \notin A_2 \\ 0 & \text{for } T = a_q b_1 \\ 0 & \text{for } T \subset N, |T| > 2, \text{ and } T \neq A_{2 \setminus 1} \cup \{b_1\} \end{cases}$$

where c is a positive number.

Then for $\forall G_3 \in \mathcal{G}_{m,n}$, we have:

$$w \cdot c_{G_3} = |A_1 \cap pa_{G_3}(b_1)| \cdot c - |pa_{G_3}(b_1) \cap A_{comp}| \cdot c$$

Now again notice that in the above equation:

- * $|A_1 \cap pa_{G_3}(b_1)| \cdot c \leq |A_1| \cdot c$, where “=” holds if and only if $A_1 \subset pa_{G_3}(b_1)$;
- * $-|pa_{G_3}(b_1) \cap A_{comp}| \cdot c \leq 0$, where “=” holds if and only if $pa_{G_3}(b_1) \cap A_{comp} = \emptyset$.

To satisfy the above two conditions, we must have $pa_{G_3}(b_1) = A_1$ or $(A_1 \cup a_q)$.

Therefore, again we have: $w \cdot c_{G_3} \leq |A_1| \cdot c$ where “=” holds if and only if $G_3 = G_1$ or G_2 .

- (2) Neither one is a subset of the other.

Define notations: $A_1 = pa_{G_1}(b_1)$, $A_2 = pa_{G_2}(b_1)$, $A_{1 \cap 2} = pa_{G_1}(b_1) \cap pa_{G_2}(b_1)$, $A_{1 \setminus 2} = pa_{G_1}(b_1) \setminus pa_{G_2}(b_1)$, $A_{2 \setminus 1} = pa_{G_2}(b_1) \setminus pa_{G_1}(b_1)$, $A_{1 \cup 2} = pa_{G_1}(b_1) \cup pa_{G_2}(b_1)$ and $A_{comp} = (A_{1 \cup 2})^c$. Notice that $A_{1 \setminus 2}$, $A_{2 \setminus 1} \neq \emptyset$, $A_{1 \cap 2}$ and A_{comp} can be \emptyset . $A_{1 \cap 2}$, $A_{1 \setminus 2}$, $A_{2 \setminus 1}$, and A_{comp} is a partition of N .

Consider a cost vector w whose coordinates match the coordinates of characteristic imsets.

– If $|A_{1\setminus 2}| > 1$ and $|A_{2\setminus 1}| > 1$. Define w as:

$$w(T) = \begin{cases} c & \text{for } T = a_i b_j, a_i \in A_{1\cap 2} \\ -c & \text{for } T = a_i b_j, a_i \notin A_{1\cap 2} \\ -2c & \text{for } T = A_{1\setminus 2} \cup A_{2\setminus 1} \cup \{b_1\} \\ (|A_{1\setminus 2}| + 1) \cdot c & \text{for } T = A_{1\setminus 2} \cup \{b_1\} \\ (|A_{2\setminus 1}| + 1) \cdot c & \text{for } T = A_{2\setminus 1} \cup \{b_1\} \\ 0 & \text{for other } T \subset N, |T| > 2 \end{cases}$$

where c is a positive number.

Now for $\forall G_3 \in \mathcal{G}_{m,n}$, we have:

$$\begin{aligned} w \cdot c_{G_3} &= |pa_{G_3}(b_1) \cap A_{1\cap 2}| \cdot c - |pa_{G_3}(b_1) \cap A_{1\setminus 2}| \cdot c \\ &\quad - |pa_{G_3}(b_1) \cap A_{2\setminus 1}| \cdot c - |pa_{G_3}(b_1) \cap A_{comp}| \cdot c \\ &\quad + (|A_{1\setminus 2}| + 1) \cdot c \cdot c_{G_3}(A_{1\setminus 2} \cup \{b_1\}) + (|A_{2\setminus 1}| + 1) \cdot c \cdot c_{G_3}(A_{2\setminus 1} \cup \{b_1\}) \\ &\quad - 2c \cdot c_{G_3}(A_{1\setminus 2} \cup A_{2\setminus 1} \cup \{b_1\}) \\ &= |pa_{G_3}(b_1) \cap A_{1\cap 2}| \cdot c \\ &\quad - |pa_{G_3}(b_1) \cap A_{1\setminus 2}| \cdot c + (|A_{1\setminus 2}| + 1) \cdot c \cdot c_{G_3}(A_{1\setminus 2} \cup \{b_1\}) \\ &\quad - |pa_{G_3}(b_1) \cap A_{2\setminus 1}| \cdot c + (|A_{2\setminus 1}| + 1) \cdot c \cdot c_{G_3}(A_{2\setminus 1} \cup \{b_1\}) \\ &\quad - 2c \cdot c_{G_3}(A_{1\setminus 2} \cup A_{2\setminus 1} \cup \{b_1\}) \\ &\quad - |pa_{G_3}(b_1) \cap A_{comp}| \cdot c \end{aligned}$$

Now notice that in the above equation:

- * $|pa_{G_3}(b_1) \cap A_{1\cap 2}| \cdot c \leq |A_{1\cap 2}| \cdot c$ where “=” holds if and only if $A_{1\cap 2} \subset pa_{G_3}(b_1)$;
- * $-|pa_{G_3}(b_1) \cap A_{1\setminus 2}| \cdot c + (|A_{1\setminus 2}| + 1) \cdot c \cdot c_{G_3}(A_{1\setminus 2} \cup \{b_1\}) \leq c$ where “=” holds if and only if $A_{1\setminus 2} \subset pa_{G_3}(b_1)$;
- * $-|pa_{G_3}(b_1) \cap A_{2\setminus 1}| \cdot c + (|A_{2\setminus 1}| + 1) \cdot c \cdot c_{G_3}(A_{2\setminus 1} \cup \{b_1\}) \leq c$ where “=” holds if and only if $A_{2\setminus 1} \subset pa_{G_3}(b_1)$;
- * $-2c \cdot c_{G_3}(A_{1\setminus 2} \cup A_{2\setminus 1} \cup \{b_1\}) \leq 0$ where “=” holds if and only if $(A_{1\setminus 2} \cup A_{2\setminus 1}) \not\subset pa_{G_3}(b_1)$;
- * $-|pa_{G_3}(b_1) \cap A_{comp}| \cdot c \leq 0$ where “=” holds if and only if $pa_{G_3}(b_1) \cap A_{comp} = \emptyset$.

The above conditions cannot be satisfied simultaneously, but after we figure out that:

- * when $pa_{G_3}(b_1) = A_{1\cap 2}$, $w \cdot c_{G_3} = |A_{1\cap 2}| \cdot c + 0 + 0 + 0 + 0 = |A_{1\cap 2}| \cdot c$;
- * when $pa_{G_3}(b_1) = A_1$, i.e. $G_3 = G_1$, $w \cdot c_{G_3} = |A_{1\cap 2}| \cdot c + c + 0 + 0 + 0 = (|A_{1\cap 2}| + 1) \cdot c$;
- * when $pa_{G_3}(b_1) = A_2$, i.e. $G_3 = G_2$, $w \cdot c_{G_3} = |A_{1\cap 2}| \cdot c + 0 + c + 0 + 0 = (|A_{1\cap 2}| + 1) \cdot c$;
- * when $pa_{G_3}(b_1) = A_{1\cup 2}$, $w \cdot c_{G_3} = |A_{1\cap 2}| \cdot c + c + c - 2c + 0 = |A_{1\cap 2}| \cdot c$,

it becomes obvious that: $w \cdot c_{G_3} \leq (|A_{1\cap 2}| + 1) \cdot c$ where “=” holds if and only if $G_3 = G_1$ or G_2 .

– If only one of $|A_{1\setminus 2}|$ and $|A_{2\setminus 1}|$ is 1. Suppose $|A_{1\setminus 2}| = 1$ and $|A_{2\setminus 1}| > 1$. Define w as:

$$w(T) = \begin{cases} c & \text{for } T = a_i b_j, a_i \in A_1 \\ -c & \text{for } T = a_i b_j, a_i \notin A_1 \\ -2c & \text{for } T = A_{1\setminus 2} \cup A_{2\setminus 1} \cup \{b_1\} \\ (|A_{2\setminus 1}| + 1) \cdot c & \text{for } T = A_{2\setminus 1} \cup \{b_1\} \\ 0 & \text{for other } T \subset N, |T| > 2 \end{cases}$$

where c is a positive number.

Now for $\forall G_3 \in \mathcal{G}_{m,n}$, we have:

$$\begin{aligned}
 w \cdot c_{G_3} &= |pa_{G_3}(b_1) \cap A_{1 \cap 2}| \cdot c + |pa_{G_3}(b_1) \cap A_{1 \setminus 2}| \cdot c \\
 &\quad - |pa_{G_3}(b_1) \cap A_{2 \setminus 1}| \cdot c - |pa_{G_3}(b_1) \cap A_{comp}| \cdot c \\
 &\quad + (|A_{2 \setminus 1}| + 1) \cdot c \cdot c_{G_3}(A_{2 \setminus 1} \cup \{b_1\}) - 2c \cdot c_{G_3}(A_{1 \setminus 2} \cup A_{2 \setminus 1} \cup \{b_1\}) \\
 &= |pa_{G_3}(b_1) \cap A_{1 \cap 2}| \cdot c \\
 &\quad + |pa_{G_3}(b_1) \cap A_{1 \setminus 2}| \cdot c \\
 &\quad - |pa_{G_3}(b_1) \cap A_{2 \setminus 1}| \cdot c + (|A_{2 \setminus 1}| + 1) \cdot c \cdot c_{G_3}(A_{2 \setminus 1} \cup \{b_1\}) \\
 &\quad - 2c \cdot c_{G_3}(A_{1 \setminus 2} \cup A_{2 \setminus 1} \cup \{b_1\}) \\
 &\quad - |pa_{G_3}(b_1) \cap A_{comp}| \cdot c
 \end{aligned}$$

Now notice that in the above equation:

- * $|pa_{G_3}(b_1) \cap A_{1 \cap 2}| \cdot c \leq |A_{1 \cap 2}| \cdot c$ where "=" holds if and only if $A_{1 \cap 2} \subset pa_{G_3}(b_1)$;
- * $|pa_{G_3}(b_1) \cap A_{1 \setminus 2}| \cdot c \leq c$ where "=" holds if and only if $A_{1 \setminus 2} \subset pa_{G_3}(b_1)$;
- * $-|pa_{G_3}(b_1) \cap A_{2 \setminus 1}| \cdot c + (|A_{2 \setminus 1}| + 1) \cdot c \cdot c_{G_3}(A_{2 \setminus 1} \cup \{b_1\}) \leq c$ where "=" holds if and only if $A_{2 \setminus 1} \subset pa_{G_3}(b_1)$;
- * $-2c \cdot c_{G_3}(A_{1 \setminus 2} \cup A_{2 \setminus 1} \cup \{b_1\}) \leq 0$ where "=" holds if and only if $(A_{1 \setminus 2} \cup A_{2 \setminus 1}) \not\subseteq pa_{G_3}(b_1)$;
- * $-|pa_{G_3}(b_1) \cap A_{comp}| \cdot c \leq 0$ where "=" holds if and only if $pa_{G_3}(b_1) \cap A_{comp} = \emptyset$.

The above conditions cannot be satisfied simultaneously, but similar with first case in (2), it can be shown that: $w \cdot c_{G_3} \leq (|A_{1 \cap 2}| + 1) \cdot c$ where "=" holds if and only if $G_3 = G_1$ or G_2 .

– If $|A_{1 \setminus 2}| = |A_{2 \setminus 1}| = 1$. Define w as:

$$w(T) = \begin{cases} c & \text{for } T = a_i b_j, a_i \in A_{1 \cup 2} \\ -c & \text{for } T = a_i b_j, a_i \notin A_{1 \cup 2} \\ -2c & \text{for } T = A_{1 \setminus 2} \cup A_{2 \setminus 1} \cup \{b_1\} \\ 0 & \text{for other } T \subset N, |T| > 2 \end{cases}$$

where c is a positive number.

Now for $\forall G_3 \in \mathcal{G}_{m,n}$, we have:

$$\begin{aligned}
 w \cdot c_{G_3} &= |pa_{G_3}(b_1) \cap A_{1 \cap 2}| \cdot c \\
 &\quad + |pa_{G_3}(b_1) \cap A_{1 \setminus 2}| \cdot c + |pa_{G_3}(b_1) \cap A_{2 \setminus 1}| \cdot c \\
 &\quad - 2c \cdot c_{G_3}(A_{1 \setminus 2} \cup A_{2 \setminus 1} \cup \{b_1\}) \\
 &\quad - |pa_{G_3}(b_1) \cap A_{comp}| \cdot c
 \end{aligned}$$

Now notice that in the above equation:

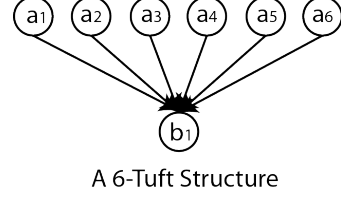
- * $|pa_{G_3}(b_1) \cap A_{1 \cap 2}| \cdot c \leq |A_{1 \cap 2}| \cdot c$ where "=" holds if and only if $A_{1 \cap 2} \subset pa_{G_3}(b_1)$;
- * $|pa_{G_3}(b_1) \cap A_{1 \setminus 2}| \cdot c \leq c$ where "=" holds if and only if $A_{1 \setminus 2} \subset pa_{G_3}(b_1)$;
- * $|pa_{G_3}(b_1) \cap A_{2 \setminus 1}| \cdot c \leq c$ where "=" holds if and only if $A_{2 \setminus 1} \subset pa_{G_3}(b_1)$;
- * $-2c \cdot c_{G_3}(A_{1 \setminus 2} \cup A_{2 \setminus 1} \cup \{b_1\}) \leq 0$ where "=" holds if and only if $(A_{1 \setminus 2} \cup A_{2 \setminus 1}) \not\subseteq pa_{G_3}(b_1)$;
- * $-|pa_{G_3}(b_1) \cap A_{comp}| \cdot c \leq 0$ where "=" holds if and only if $pa_{G_3}(b_1) \cap A_{comp} = \emptyset$.

The above conditions cannot be satisfied simultaneously, but similar with first case in (2), it can be shown that: $w \cdot c_{G_3} \leq (|A_{1 \cap 2}| + 1) \cdot c$ where "=" holds if and only if $G_3 = G_1$ or G_2 .

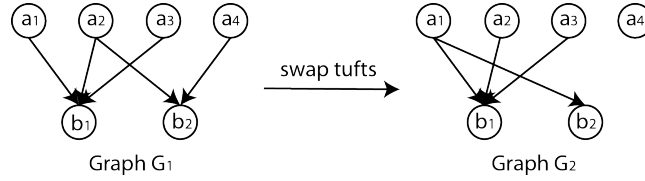
□

Remark 4.6. We can understand Lemma 4.5 better after we introduce a new concept "tuft".

- A structure is called **k-tuft** if it includes $(k + 1)$ nodes such that there exists one node (the **offspring**) that all other nodes are its parents. There are two important special cases: an 1-tuft is just an edge and a 2-tuft is just an immortality. An example of 6-tuft is given in the right hand side graph.



- Fix m and n . We say a change from one graph G_1 to another one G_2 is **swapping turves** if: G_2 can be obtained from G_1 by removing a tuft and adding another tuft, where these two turves share the same offspring but disjoint parents. An example is given as below: G_2 can be obtained from G_1 by removing a 2-tuft and add an 1-tuft, where the two turves share the same offspring b_2 .



- Lemma 4.5 gives us a principal idea of finding neighbors for a graph: add or remove a tuft, or swap turves once. We are going to prove this in general case in Theorem 4.7.

Theorem 4.7. *Notation same as Definition 2.1. Fix m and n . Two graphs, $G_1, G_2 \in \mathcal{G}_{m,n}$ are neighbors if and only if $\exists b_i \in B$ such that $pa_{G_1}(b_i) \neq pa_{G_2}(b_i)$ and $pa_{G_1}(b_j) = pa_{G_2}(b_j)$, for $\forall b_j \in B$ and $b_j \neq b_i$, i.e., G_2 can be obtained from G_1 by removing or adding a tuft or swapping turves.*

Proof. We will prove “if” and “only if” separately.

(1) Prove “if” part.

Suppose we have two graphs, $G_1, G_2 \in \mathcal{G}_{m,n}$, and $\exists b_i \in B$ such that $pa_{G_1}(b_i) \neq pa_{G_2}(b_i)$ and $pa_{G_1}(b_j) = pa_{G_2}(b_j)$, for $\forall b_j \in B$ and $b_j \neq b_i$. We need to prove G_1 and G_2 are neighbors.

Consider an arbitrary graph $G_3 \in \mathcal{G}_{m,n}$. We want to prove that \exists a cost vector w such that $w \cdot c_{G_1} = w \cdot c_{G_2} \geq w \cdot c_{G_3}$ where “=” holds if and only if $G_3 = G_1$ or G_2 .

Define the following graphs (see Remark 4.8 for example):

- ★ $G'_1, G'_2, G'_3 \in \mathcal{G}_{m,1}$ with symptom $B_{m,1} = \{b_i\}$ such that $pa_{G'_1}(b_i) = pa_{G_1}(b_i)$, $pa_{G'_2}(b_i) = pa_{G_2}(b_i)$ and $pa_{G'_3}(b_i) = pa_{G_3}(b_i)$;
- ★ $G_0, G''_3 \in \mathcal{G}_{m,(n-1)}$ with symptoms $B_{m,(n-1)} = B \setminus \{b_i\}$ such that $pa_{G_0}(b_j) = pa_{G_1}(b_j) = pa_{G_2}(b_j)$ and $pa_{G''_3}(b_j) = pa_{G_3}(b_j)$, $\forall b_j \in B_{m,(n-1)}$.

Notice that the connections of the characteristic imsets of these graphs are very simple. By Remark 3.5, after moving the coordinates properly, we can write the characteristic imsets of G_1, G_2 and G_3 in the form of:

$$\begin{aligned} c_{G_1} &= (c_{G'_1} \quad c_{G_0}) \\ c_{G_2} &= (c_{G'_2} \quad c_{G_0}) \\ c_{G_3} &= (c_{G'_3} \quad c_{G''_3}) \end{aligned}$$

- As proved in Lemma 4.5, G'_1 and G'_2 are neighbors. This means that \exists related cost vector w_1 such that $w_1 \cdot c_{G'_1} = w_1 \cdot c_{G'_2} \geq w_1 \cdot c_{G'_3}$ for $\forall G'_3 \in \mathcal{G}_{m,1}$, where “=” holds if and only if $G'_3 = G'_1$ or G'_2 .
- Because c_{G_0} is a vertex of the characteristic inset polytope related to $\mathcal{G}_{m,(n-1)}$, we can find a related cost vector w_2 such that $w_2 \cdot c_{G_0} \geq w_2 \cdot c_{G''_3}$ for $\forall G''_3 \in \mathcal{G}_{m,(n-1)}$, where “=” holds if and only if $G''_3 = G_0$.

Now let $w = (w_1 \ w_2)$ with the new permutation of coordinates. We have:

$$\begin{aligned} w \cdot c_{G_1} &= w_1 \cdot c_{G'_1} + w_2 \cdot c_{G_0} \\ &= w_1 \cdot c_{G'_2} + w_2 \cdot c_{G_0} = w \cdot c_{G_2} \\ &\geq w_1 \cdot c_{G'_3} + w_2 \cdot c_{G''_3} = w \cdot c_{G_3} \end{aligned}$$

where “=” holds if and only if i) $G'_3 = G'_1$ or G'_2 , and ii) $G''_3 = G_0$, i.e. $G_3 = G_1$ or G_2 .

(2) Prove “only if” part.

Suppose we have two graphs, $G_1, G_2 \in \mathcal{G}_{m,n}$, which are neighbors. i.e. \exists a cost vector w such that $w \cdot c_{G_1} = w \cdot c_{G_2} > w \cdot c_G$ for $\forall G \in \mathcal{G}_{m,n}$, $G \neq G_1, G_2$. We are going to prove by contradiction.

Suppose $\exists b_i, b_j \in B$ distinct, $pa_{G_1}(b_i) \neq pa_{G_2}(b_i)$ and $pa_{G_1}(b_j) \neq pa_{G_2}(b_j)$.

Define the following graphs (see Remark 4.8 for example):

- * $G'_1, G'_2 \in \mathcal{G}_{m,1}$ with symptom $B_{m,1} = \{b_i\}$ such that $pa_{G'_1}(b_i) = pa_{G_1}(b_i)$ and $pa_{G'_2}(b_i) = pa_{G_2}(b_i)$;
- * $G''_1, G''_2 \in \mathcal{G}_{m,1}$ with symptom $B_{m,1} = \{b_j\}$ such that $pa_{G''_1}(b_j) = pa_{G_1}(b_j)$ and $pa_{G''_2}(b_j) = pa_{G_2}(b_j)$;
- * $G'''_1, G'''_2 \in \mathcal{G}_{m,(n-2)}$ with symptoms $B_{m,(n-2)} = B \setminus \{b_i, b_j\}$ such that $pa_{G'''_1}(b_k) = pa_{G_1}(b_k)$ and $pa_{G'''_2}(b_k) = pa_{G_2}(b_k)$, $\forall b_k \in B_{m,(n-2)}$;
- * $G_3 \in \mathcal{G}_{m,n}$ is all same with G_1 but $pa_{G_3}(b_i) = pa_{G_2}(b_i)$;
- * $G_4 \in \mathcal{G}_{m,n}$ is all same with G_1 but $pa_{G_4}(b_j) = pa_{G_2}(b_j)$;
- * $G_5 \in \mathcal{G}_{m,n}$ is all same with G_2 but $pa_{G_5}(b_i) = pa_{G_1}(b_i)$ and $pa_{G_5}(b_j) = pa_{G_1}(b_j)$, notice that G_5 might be same with G_1 .

Similar with (1), after moving the coordinates properly, we can write the characteristic insets of G_1, G_2, G_3, G_4 and G_5 in the following form:

$$\begin{aligned} c_{G_1} &= (c_{G'_1} \ c_{G''_1} \ c_{G'''_1}) \\ c_{G_2} &= (c_{G'_2} \ c_{G''_2} \ c_{G'''_2}) \\ c_{G_3} &= (c_{G'_2} \ c_{G''_1} \ c_{G'''_1}) \\ c_{G_4} &= (c_{G'_1} \ c_{G''_2} \ c_{G'''_1}) \\ c_{G_5} &= (c_{G'_1} \ c_{G''_1} \ c_{G'''_2}) \end{aligned}$$

Now use the same permutation of coordinates in w and do the same partition, we can write w in the form of $w = (w_1 \ w_2 \ w_3)$. By the assumption we indicated at the beginning of this part, we have:

- * It is obvious that $G_3 \neq G_1$ or G_2 . So:

$$\begin{aligned} w \cdot c_{G_1} &= w_1 \cdot c_{G'_1} + w_2 \cdot c_{G''_1} + w_3 \cdot c_{G'''_1} \\ &> w \cdot c_{G_3} = w_1 \cdot c_{G'_2} + w_2 \cdot c_{G''_1} + w_3 \cdot c_{G'''_1} \\ \implies w_1 \cdot c_{G'_1} &> w_1 \cdot c_{G'_2} \end{aligned}$$

- * It is obvious that $G_4 \neq G_1$ or G_2 . So:

$$\begin{aligned} w \cdot c_{G_1} &= w_1 \cdot c_{G'_1} + w_2 \cdot c_{G''_1} + w_3 \cdot c_{G'''_1} \\ &> w \cdot c_{G_4} = w_1 \cdot c_{G'_1} + w_2 \cdot c_{G''_2} + w_3 \cdot c_{G'''_1} \\ \implies w_2 \cdot c_{G''_1} &> w_2 \cdot c_{G''_2} \end{aligned}$$

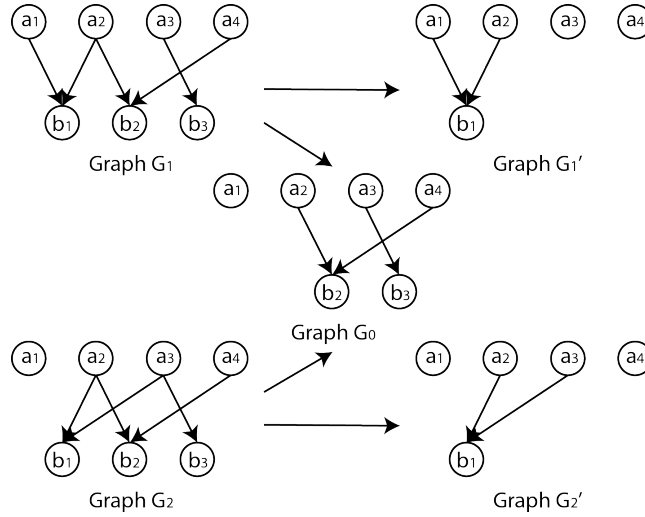
Then we have:

$$\begin{aligned}
 w \cdot c_{G_2} &= w_1 \cdot c_{G'_2} + w_2 \cdot c_{G''_2} + w_3 \cdot c_{G'''_2} \\
 &< w_1 \cdot c_{G'_1} + w_2 \cdot c_{G''_1} + w_3 \cdot c_{G'''_1} = w \cdot c_{G_5} \\
 \implies w \cdot c_{G_2} &< w \cdot c_{G_5}
 \end{aligned}$$

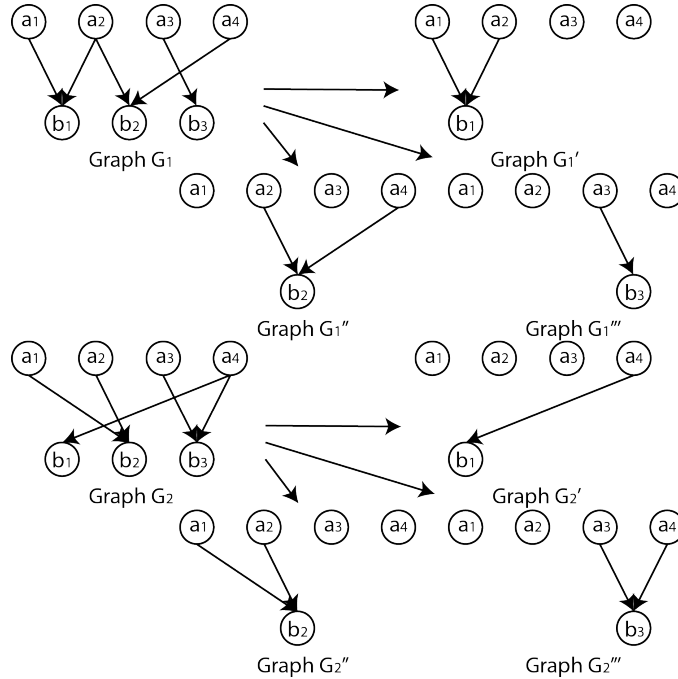
which is a contradiction with our assumption. Therefore G_1 and G_2 cannot be neighbors. \square

Remark 4.8. In the two parts of proof of Theorem 4.7, it might be more intuitive to see how we define the new graphs using examples.

- Part (1), prove “if” condition. Let $m = 4$ and $n = 3$. In the example below, $b_i = b_1$.



- Part (2), prove “only if” condition. Let $m = 4$ and $n = 3$. In the example below, $b_i = b_1$ and $b_j = b_2$.



Theorem 4.9. *Notation same as Definition 2.1. Fix m and n . For $\forall G \in \mathcal{G}_{m,n}$, G has $n \cdot (2^m - 1)$ many neighbors.*

Proof. By Theorem 4.7, for $\forall H \in \mathcal{G}_{m,n}$, G and H are neighbors if and only if: $\exists b_i \in B$ such that $pa_G(b_i) \neq pa_H(b_i)$ and $pa_G(b_j) = pa_H(b_j)$, for $\forall b_j \in B$ and $b_j \neq b_i$.

For a fixed b_i , define graphs:

- ★ $G', H' \in \mathcal{G}_{m,1}$ with symptom $B_{m,1} = \{b_i\}$ such that $pa_{G'}(b_i) = pa_G(b_i)$ and $pa_{H'}(b_i) = pa_H(b_i)$;
- ★ $G'', H'' \in \mathcal{G}_{m,(n-1)}$ with symptoms $B_{m,(n-1)} = B \setminus \{b_i\}$ such that $pa_{G''}(b_j) = pa_G(b_j)$ and $pa_{H''}(b_j) = pa_H(b_j)$, $\forall b_j \in B_{m,(n-1)}$.

Now suppose G and H are neighbors and $G' \neq H'$, then $G'' = H''$. By Proposition 3.6, there are 2^m graphs in $\mathcal{G}_{m,1}$. This means that there are $2^m - 1$ possible different H' 's, and each one relate to a different neighbor of G .

We can use the same strategy for each symptom, i.e. we can find $2^m - 1$ neighbors for each $b_i \in B$. It is easy to see that these neighbors are all distinct:

- ★ H_1, H_2 are all same with G but $pa_G(b_i) \neq pa_{H_1}(b_i)$ and $pa_G(b_j) \neq pa_{H_2}(b_j)$, where $b_i, b_j \in B$ are distinct. Then $pa_{H_2}(b_i) = pa_G(b_i) \neq pa_{H_1}(b_i)$, so H_1 and H_2 are different.

Therefore the total number of neighbors for G is: $n \cdot (2^m - 1)$. \square

Remark 4.10. Recall that we proved in Theorem 4.1, for fixed m and n , the dimension of the characteristic imset polytope is $n \cdot (2^m - 1)$. Thus Theorem 4.9 implies that the number of neighbors for each vertex equals to the dimension, i.e. the characteristic imset polytope for Bipartite graphs is a **simple polytope**. In 2000, V. Kaibel and M. Wolff proved that a zero-one polytope is simple if and only if it equals to a direct product of zero-one simplices [5]. But here, we are going to prove a even stronger result.

Theorem 4.11. *Notation same with Definition 2.1. $\mathbf{P}_{m,n}$ is the direct product of n many Δ_{2^m-1} , i.e.*

$$\mathbf{P}_{m,n} = \underbrace{\Delta_{2^m-1} \times \Delta_{2^m-1} \times \cdots \times \Delta_{2^m-1}}_{n \text{ many}}.$$

And the i_{th} simplex is $\mathbf{P}_{m,1}$ with the same diseases A and one symptom $\{b_i\}$.

Proof. Fix m , we are going to prove the equality holds using induction on n .

- $n = 1$. Proved in Theorem 4.3;
- Fix $q \in \mathbb{Z}^+$. Suppose the equality holds for $\mathbf{P}_{m,n}$, $\forall n < q$, and we need to prove that it also holds for $\mathbf{P}_{m,q}$. Recall that for $\mathcal{G}_{m,q}$, we have notation for all symptoms: $B = \{b_1, b_2, \dots, b_q\}$.

First, we want to prove: $\mathbf{P}_{m,n} \subseteq \mathbf{P}_{m,q-1} \times \mathbf{P}_{m,1}$.

Similar with the proof of Theorem 4.7, for $\forall G \in \mathcal{G}_{m,n}$, we can define graphs:

- ★ $G' \in \mathcal{G}_{m,(q-1)}$ with symptoms $B_{m,(q-1)} = B \setminus \{b_q\}$ such that $pa_{G'}(b_i) = pa_G(b_i)$, $\forall b_i \in B_{m,(q-1)}$. This implies $c_{G'} \in \mathbf{P}_{m,q-1}$;
- ★ $G'' \in \mathcal{G}_{m,1}$ with symptom $B_{m,1} = \{b_q\}$ such that $pa_{G''}(b_q) = pa_G(b_q)$. This implies $c_{G''} \in \mathbf{P}_{m,1}$.

Again, with a proper permutation of coordinates, we can write c_G in form of:

$$c_G = (c_{G'} \ c_{G''}).$$

Now because the set of vertices of $\mathbf{P}_{m,q}$ is $\{c_G : G \in \mathcal{G}_{m,q}\}$, so for $\forall x \in \mathbf{P}_{m,q}$, with the same permutation of coordinates, we have:

$$x = \sum_{G \in \mathcal{G}_{m,q}} \alpha_G c_G = \left(\sum_{G \in \mathcal{G}_{m,q}} \alpha_G c_{G'} \ , \ \sum_{G \in \mathcal{G}_{m,q}} \alpha_G c_{G''} \right),$$

where $0 \leq \alpha_G \leq 1$, $\forall G \in \mathcal{G}_{m,q}$ and $\sum_{G \in \mathcal{G}_{m,q}} \alpha_G = 1$.

Notice that $\sum_{G \in \mathcal{G}_{m,q}} \alpha_G c_{G'} \in \mathbf{P}_{m,q-1}$ and $\sum_{G \in \mathcal{G}_{m,q}} \alpha_G c_{G''} \in \mathbf{P}_{m,1}$, the above equality implies $x \in \mathbf{P}_{m,q-1} \times \mathbf{P}_{m,1}$. Hence:

$$\mathbf{P}_{m,q} \subseteq \mathbf{P}_{m,q-1} \times \mathbf{P}_{m,1}.$$

Second, we want to prove: $\mathbf{P}_{m,q-1} \times \mathbf{P}_{m,1} \subseteq \mathbf{P}_{m,n}$.

Let $\mathcal{G}_{m,q-1}$ has symptoms $B_{m,(q-1)} = B \setminus \{b_q\}$ and $\mathcal{G}_{m,1}$ has symptom $B_{m,1} = \{b_q\}$. For $\forall G' \in \mathcal{G}_{m,(q-1)}$ and $G'' \in \mathcal{G}_{m,1}$, we can define $G \in \mathcal{G}_{m,q}$ such that $pa_G(b_i) = pa_{G'}(b_i)$, $\forall b_i \in B_{m,(q-1)}$, and $pa_G(b_q) = pa_{G''}(b_q)$. We can write c_G in form of $c_G = (c_{G'} \ c_{G''})$.

Now for $\forall x \in \mathbf{P}_{m,q-1} \times \mathbf{P}_{m,1}$, by definition of direct product, it can be written as:

$$\begin{aligned} x &= \left(\sum_{G' \in \mathcal{G}_{m,q-1}} \beta_{G'} c_{G'}, \sum_{G'' \in \mathcal{G}_{m,1}} \gamma_{G''} c_{G''} \right) = \sum_{G' \in \mathcal{G}_{m,q-1}} \sum_{G'' \in \mathcal{G}_{m,1}} \beta_{G'} \gamma_{G''} (c_{G'}, c_{G''}) \\ &= \sum_{G' \in \mathcal{G}_{m,q-1}} \sum_{G'' \in \mathcal{G}_{m,1}} (\beta_{G'} \gamma_{G''}) c_G, \end{aligned}$$

where $0 \leq \beta_{G'}, \gamma_{G''} \leq 1$, $\forall G' \in \mathcal{G}_{m,q-1}$, $G'' \in \mathcal{G}_{m,1}$, and $\sum_{G' \in \mathcal{G}_{m,q-1}} \beta_{G'} = 1$, $\sum_{G'' \in \mathcal{G}_{m,1}} \gamma_{G''} = 1$.

Notice that

$$\sum_{G' \in \mathcal{G}_{m,q-1}} \sum_{G'' \in \mathcal{G}_{m,1}} (\beta_{G'} \gamma_{G''}) = \sum_{G' \in \mathcal{G}_{m,q-1}} \beta_{G'} \left(\sum_{G'' \in \mathcal{G}_{m,1}} \gamma_{G''} \right) = \sum_{G' \in \mathcal{G}_{m,q-1}} \beta_{G'} = 1.$$

This leads to $x \in \mathbf{P}_{m,q}$. Hence:

$$\mathbf{P}_{m,q-1} \times \mathbf{P}_{m,1} \subseteq \mathbf{P}_{m,n}.$$

Now using the assumption we made before, we can finish the proof because:

$$\mathbf{P}_{m,q} = \mathbf{P}_{m,q-1} \times \mathbf{P}_{m,1} = \underbrace{\Delta_{2^m-1} \times \dots \times \Delta_{2^m-1}}_{q-1 \text{ many}} \times \Delta_{2^m-1} = \underbrace{\Delta_{2^m-1} \times \dots \times \Delta_{2^m-1}}_{q \text{ many}}.$$

□

4.2. Facets of $\mathbf{P}_{m,n}$.

Recall that there is a very famous theorem [14] about facets for direct product of simplices.

Lemma 4.12. [14] *Denote \mathbf{P} as the direct product of simplices $\Delta_{\alpha_1}, \dots, \Delta_{\alpha_k}$. Then every facet of \mathbf{P} has the form of $\Delta_{\alpha_1} \times \dots \times \Delta_{\alpha_{i-1}} \times F_{\alpha_i} \times \Delta_{\alpha_{i+1}} \times \dots \times \Delta_{\alpha_k}$, where F_{α_i} is a facet of Δ_{α_i} .*

Remark 4.13. This Lemma implies that if we want to study the facets of a direct product of simplices, we can simplify the problem as studying facets of each simplex. Now because we already proved in Theorem 4.11 that $\mathbf{P}_{m,n}$ is a direct product of n many $\mathbf{P}_{m,1}$, our problem is simplified as studying the facets of $\mathbf{P}_{m,1}$. Therefore, in the following part of this subsection we will assume $B = \{b_1\}$.

Assume $A = \{a_1, \dots, a_m\}$ and $B = \{b_1\}$. As shown in Section 3, the vertices of $\mathbf{P}_{m,1}$ has at most $2^m - 1$ many non-zero coordinates. We can remark the coordinates as: $x = \{x_s, s \subseteq A, s \neq \emptyset\}$, here each x_s corresponds to one coordinate in the characteristic imset $c_G(s \cup \{b_1\})$. Now suppose $A_m x \leq b_m$ is the system of inequalities that defines $\mathbf{P}_{m,1}$, and we can define a $2^m \times 2^m$ matrix: $D_m = [b_m] - A_m$. We can denote the elements in D_m as $(d_{st})_{s \subseteq A, t \subseteq A}$ such that we can rewrite the system of inequalities as: $d_{s\emptyset} + \sum_{t \subseteq A, t \neq \emptyset} d_{st} x_t \geq 0$, $s \subseteq A$. Then we have the expression of 2^m facets of $\mathbf{P}_{m,1}$ as following:

$$F_s = \mathbf{P}_{m,1} \cap \{x : d_{s\emptyset} + \sum_{t \subseteq A, t \neq \emptyset} d_{st} x_t = 0\}, \quad s \subseteq A,$$

where the elements d_{st} , $s, t \subseteq A$ can be obtained using the following theorem.

Theorem 4.14. *The elements in matrix D_m can be computed as following:*

- $d_{st} \neq 0$ if and only if $s \subseteq t$;
- if $s \subseteq t$, then $d_{st} = (-1)^{|t|-|s|}$.

This implies that $\mathbf{P}_{m,1}$ has 2^m facets:

$$F_s = \mathbf{P}_{m,1} \cap \{x : d_{s\emptyset} + \sum_{t \subseteq A, t \neq \emptyset} d_{st} x_t = 0\}, \quad s \subseteq A.$$

What's more, for $\forall s \subseteq A$, all vertices of $\mathbf{P}_{m,1}$ are on F_s only except c_{G_s} where $pa_{G_s}(b_1) = s$.

Proof. Let $x_\emptyset \equiv 1$. For $\forall s \subseteq A$, let $d_s \cdot$ be the row of D_m which has index s , and G_s be the graph in $\mathbf{P}_{m,1}$ for which $pa_{G_s}(b_1) = s$. Now we can rewrite the system of inequalities as:

$$\sum_{t \subseteq A} d_{st} x_t = d_s \cdot (1 \ x)^T \geq 0, \quad \text{for } \forall s \subseteq A$$

We are going to prove that for $\forall s \subseteq A$, we can find $2^m - 1$ vertices on F_s which are linearly independent, and this implies that F_s is a facet of $\mathbf{P}_{m,1}$. Actually we will prove that: $\{c_{G_{s'}}, s' \subseteq A, s' \neq s\} \subset F_s$ and $c_{G_s} \notin F_s$, i.e. $d_s \cdot (1 \ c_{G_{s'}})^T = 0$, $\forall s' \subseteq A, s' \neq s$ and $d_s \cdot (1 \ c_{G_s})^T > 0$. Notice that for $\forall t \subseteq A$, $c_{G_{s'}}(t \cup \{b_1\}) \neq 0$ if and only if $t \subseteq pa_{c_{G_{s'}}}(b_1) = s'$, and $d_{st} \neq 0$ if and only if $s \subseteq t$. Now we have:

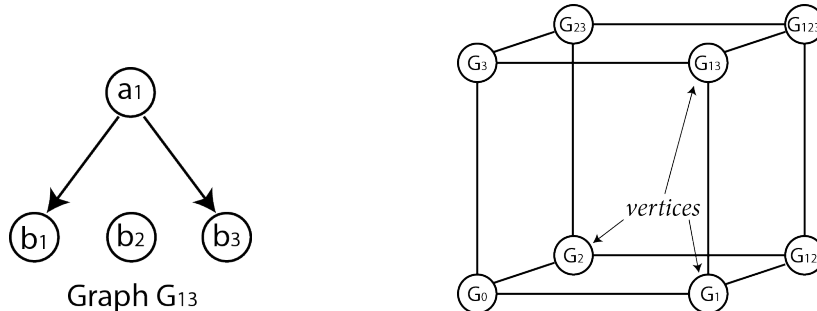
$$d_s \cdot (1 \ c_{G_{s'}})^T = d_{s\emptyset} + \sum_{t \subseteq A, t \neq \emptyset} d_{st} c_{G_{s'}}(t \cup \{b_1\}) = d_{s\emptyset} + \sum_{s \subseteq t \subseteq s', t \neq \emptyset} d_{st} = \sum_{s \subseteq t \subseteq s'} d_{st}.$$

- If $s = s'$, then $d_s \cdot (1 \ c_{G_{s'}})^T = d_{ss} = 1 > 0$;
- If $s \subsetneq s'$, then $d_s \cdot (1 \ c_{G_{s'}})^T = \sum_{s \subseteq t \subseteq s'} (-1)^{|t|-|s|} = \sum_{t' \subseteq s' \setminus s} (-1)^{|t'|} = 0$;
- If $s \not\subseteq s'$, then $d_s \cdot (1 \ c_{G_{s'}})^T = 0$.

□

5. EXAMPLES

Example 5.1 (Only One Disease). For $m = 1$, by Proposition 3.2, only $c_G(a_1 b_j)$, $b_j \in \{1, \dots, n\}$, can be 1. So consider all combination of existence of edge $(a_1 b_j)$, $b_j \in \{1, \dots, n\}$, it is easy to see that the polytope of all possible characteristic imset is **the n -dimensional unit hypercube**. A simple example of $n = 3$ is given here.



We can list all possible characteristic imsets in a table.

$$\begin{pmatrix} c_{G_0} \\ c_{G_1} \\ c_{G_2} \\ c_{G_3} \\ c_{G_{12}} \\ c_{G_{23}} \\ c_{G_{13}} \\ c_{G_{123}} \end{pmatrix} = T \begin{pmatrix} a_1b_1 & a_1b_2 & a_1b_3 & b_1b_2 & b_2b_3 & b_1b_3 & a_1b_1b_2 & a_1b_2b_3 & a_1b_1b_3 & b_1b_2b_3 & a_1b_1b_2b_3 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Example 5.2 ($m = 2, n = 2$). Use Proposition 3.3, we can easily figure out all possible characteristic imsets. An example for encoding the subscript is given by Figure 2.

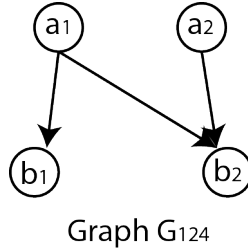
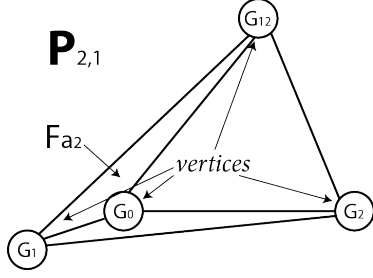


FIGURE 2. Encode edges a_1b_1 as 1, a_1b_2 as 2, a_2b_1 as 3, a_2b_2 as 4. Graph G_{124} is shown as above.

$$\begin{pmatrix} c_{G_0} \\ c_{G_1} \\ c_{G_2} \\ c_{G_3} \\ c_{G_4} \\ c_{G_{12}} \\ c_{G_{13}} \\ c_{G_{14}} \\ c_{G_{23}} \\ c_{G_{24}} \\ c_{G_{34}} \\ c_{G_{123}} \\ c_{G_{134}} \\ c_{G_{124}} \\ c_{G_{234}} \\ c_{G_{1234}} \end{pmatrix} = T \begin{pmatrix} a_1b_1 & a_1b_2 & a_2b_1 & a_2b_2 & b_1b_2 & a_1a_2b_1 & a_1a_2b_2 & a_1b_1b_2 & a_2b_1b_2 & a_1a_2b_1b_2 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$

Example 5.3 (Facets of a simple example). Fix $m = 2$ and $n = 1$.



The matrix $D_2 = [b_2] - A_2$:

$$D_2 = \begin{array}{c|cccc} s \backslash t & \emptyset & a_1 & a_2 & a_1 a_2 \\ \hline \emptyset & 1 & -1 & -1 & 1 \\ a_1 & 0 & 1 & 0 & -1 \\ a_2 & 0 & 0 & 1 & -1 \\ a_1 a_2 & 0 & 0 & 0 & 1 \end{array}$$

The system of inequalities which defines $\mathbf{P}_{2,1}$:

$$\begin{pmatrix} c_{G_0} \\ c_{G_1} \\ c_{G_2} \\ c_{G_{12}} \end{pmatrix} = \begin{array}{c|ccc} T & a_1 b_1 & a_2 b_1 & a_1 a_2 b_1 \\ \hline \begin{pmatrix} c_{G_0} \\ c_{G_1} \\ c_{G_2} \\ c_{G_{12}} \end{pmatrix} & \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \end{array} \quad \begin{array}{c|cccc} s \backslash t & \emptyset & a_1 & a_2 & a_1 a_2 \\ \hline \emptyset & 1 & -x_{a_1} & -x_{a_2} & +x_{a_1 a_2} \geq 0 \\ a_1 & & x_{a_1} & & -x_{a_1 a_2} \geq 0 \\ a_2 & & & x_{a_2} & -x_{a_1 a_2} \geq 0 \\ a_1 a_2 & & & & x_{a_1 a_2} \geq 0 \end{array}$$

The facet F_{a_2} is determined by vertices c_{G_0} , c_{G_1} and $c_{G_{12}}$.

6. DISCUSSION

In this paper we focus on the diagnosis models and their characteristic imsets. We have shown that the dimension of the characteristic imset polytope for the diagnosis models with fixed m and n is $n \cdot (2^m - 1)$, the explicit description of all edges of the polytope in terms of graphs, and the characteristic imset polytope for the diagnosis models with fixed m and n are the direct product of n many $(2^m - 1)$ dimensional simplices. In the following subsection we have listed several questions which we are working on.

6.1. Future work. There are several things we would like to work on the characteristic imset polytopes for diagnosis models $\mathbf{P}_{m,n}$. First, we would like to find all facets of $\mathbf{P}_{m,n}$ for any $m, n \geq 1$. Then in order to analyze sensitivity of the quality criteria and data, we would like to compute the *normal fan* of $\mathbf{P}_{m,n}$. Consider a vertex v of $\mathbf{P}_{m,n}$. A *normal cone* at v is a cone generated by the normal vectors of all facets which contain v . Note that the normal cone at v is the set of all cost functions c which give the vertex v as the optima solution for the linear programming problem,

$$\max c \cdot x \text{ such that } x \in \mathbf{P}_{m,n}.$$

The normal fan is the union of normal cones for all vertices of the polytope. It is also interesting to consider the mis-specification problem, that is, the behavior of optimal solution when the assumption of diagnosis model is wrong.

Also we would like to study the characteristic imset polytope for trees with N nodes.

7. ACKNOWLEDGMENT

The authors would like to thank M. Studený and Bernd Sturmfels for their useful advice.

REFERENCES

- [1] H. Akaike. Information theory and an extension of the maximum likelihood principle. In B. Petrox & F. Caski (Eds.), *Proceedings of the second international symposium on information theory*, pages 267–281, 1973.
- [2] S. A. Andersson, D. Madigan, and M. D. Perlman. A characterization of markov equivalence classes for acyclic digraphs. *Annals of Statistics*, 25:505–541, 1997.
- [3] N. Friedman, M. Linial, I. Nachman, and D. Pe’er. Using bayesian networks to analyze expression data. *Journal of Computational Biology*, 7:601–620, 2000.

- [4] X. Jiang, R.E. Neapolitan, M.M. Barmada, and S. Visweswaran. Learning genetic epistasis using bayesian network scoring criteria. *BMC Bioinformatics*, 12(89), 2011.
- [5] V. Kaibel and M. Wolff. Simple 0/1-polytope. *Europ. J. Combinatorics*, 21:139–144, 2000.
- [6] S. L. Lauritzen. *Graphical Models*. Clarendon Press, 1996.
- [7] P.J.F. Lucas. Bayesian model-based diagnosis. *International Journal of Approximate Reasoning*, 27(2):99–119, 2001.
- [8] G. Schwarz. Estimating the dimension of a model. *Annals of Statistics*, 6:461–464, 1978.
- [9] M. A. Shwe, D. E. Heckerman, M. Henrion, H. P. Lehmann, and G. F. Cooper. Probabilistic diagnosis using a reformulation of the internist-1/qmr knowledge base: I. the probabilistic model and inference algorithms. *Methods of Information in Medicine*, 30:241–255, 1991.
- [10] M. Studený. *Probabilistic Conditional Independence Structures*. Springer Verlag, 2005.
- [11] M. Studený, R. Hemmecke, and S. Lindner. Characteristic imset: a simple algebraic representative of a bayesian network structure. In *Proceedings of the 5th European Workshop on Probabilistic Graphical Models*, pages 257–264, 2010.
- [12] M. Studený, J. Vomlel, and R. Hemmecke. A geometric view on learning bayesian network structures. *International Journal of Approximate Reasoning*, 51(5):573–586, 2010.
- [13] J. Uebersax. Pgenetic counseling and cancer risk modeling: An application of bayes nets. marbella. *Spain: Ravenpack International*, 2004.
- [14] G. Ziegler. *Lectures on Polytopes*. Springer Verlag, New York, New York, 1994.